Introduction to the BNL-711 PCIe Card

Kai Chen

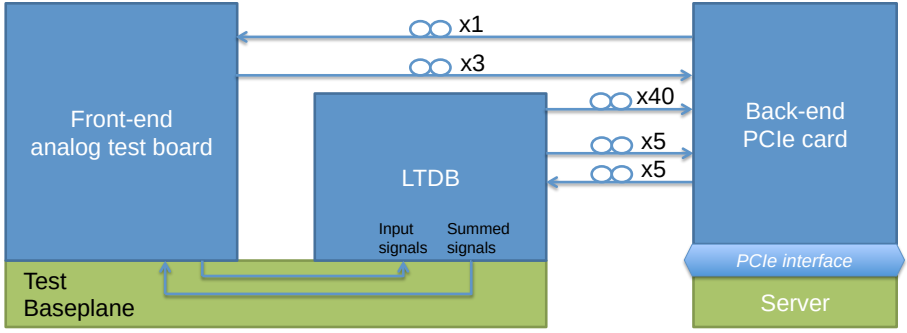
sPHENIX TPC DAM Development Meeting
December 19, 2016

Outline

Background of the BNL-711 development

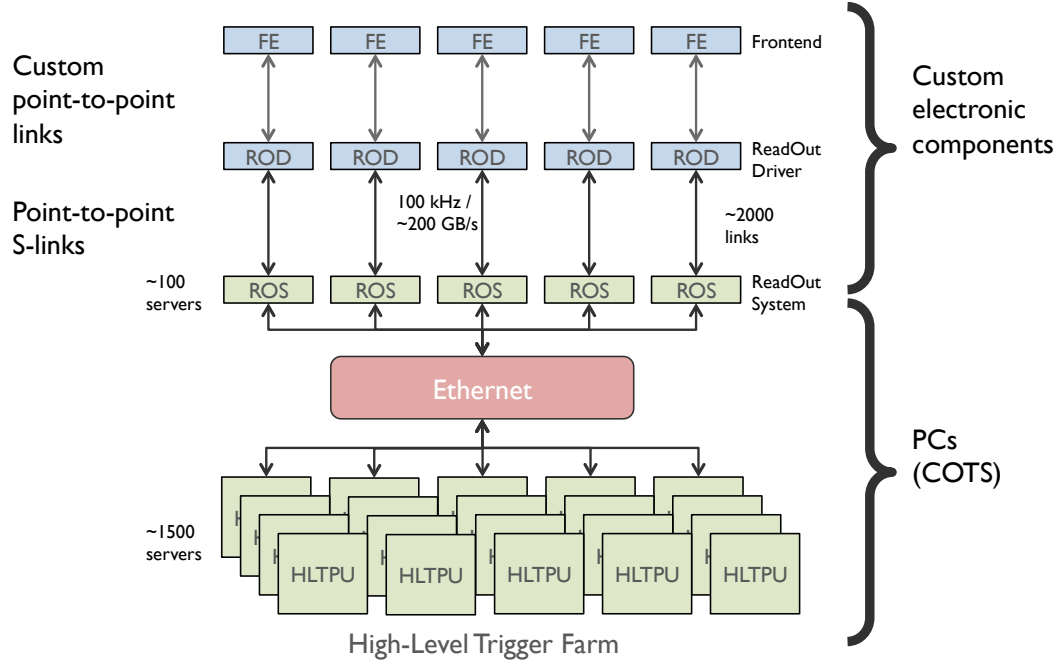
- BNL-711 is originally designed for the ATLAS LAr Phase-I upgrade LTDB (LAr Trigger Digitizer Board) test setup.
- Meanwhile BNL is involved in the ATLAS FELIX project. There is no proper commercial board available for FELIX.
- Since the BNL-711 meets FELIX requirement, it becomes the baseline choice of FELIX prototype.

Back-end of LTDB test setup



- LTDB is a board to process and digitize the 320-channel super cell LAr signals, then transmit it to the back-end.
- 5× 4.8 Gb/s GBT (GigaBit Transceiver protocol) links for slow-control and clock distribution, it communicate with GBTx ASIC on LTDB.
- 40× 5.12 Gb/s RX Data links in LTDB specialized protocol.
- 3× links for slow-control and data readout with front-end analog test board.
- DDR4 are used to buffer the 320-ch data at the same time.

ATLAS DAQ Today

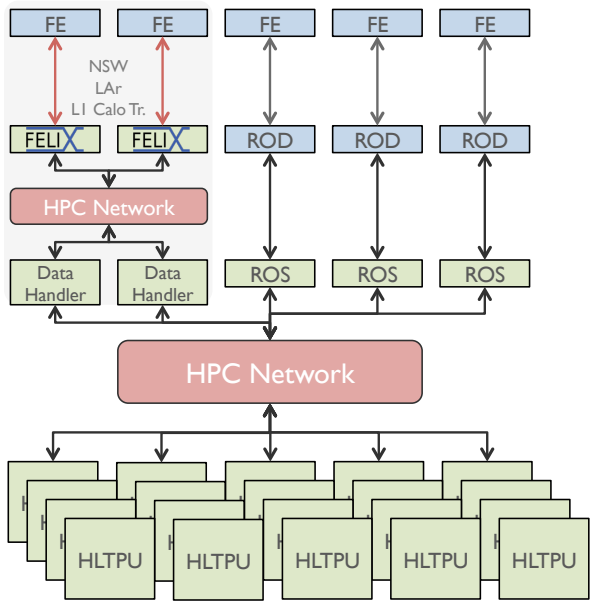


ATLAS DAQ for Phase-I

Versatile Link, GBT (GigaBit Transceiver)

PCs

40 Gb Ethernet, Infiniband



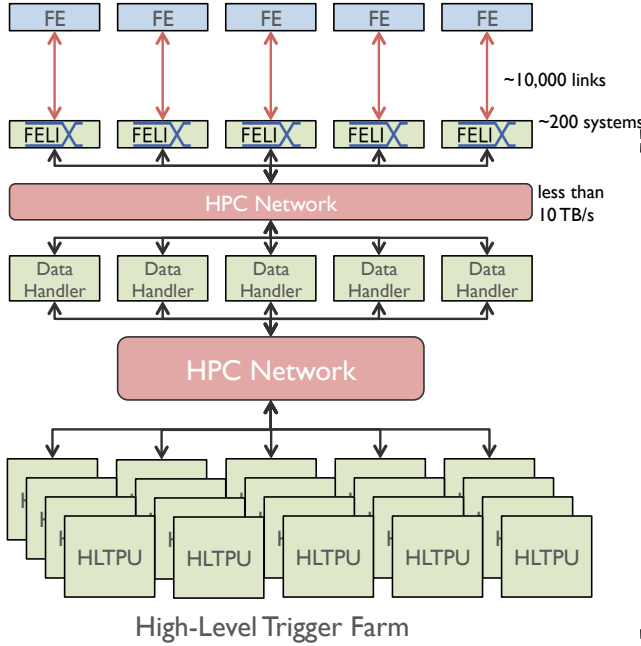
Custom electronic components including FELIX cards

PCs (COTS)

ATLAS DAQ for HL-LHC

Versatile Link, GBT, LpGBT (Low power GBT)

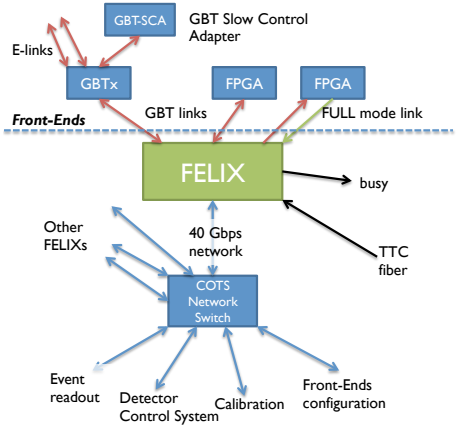
COTS network technology



Custom electronic components including FELIX cards

PCs (COTS)

FELIX functionality



- Normal GBT mode: 3.2 Gbps payload, with FEC (forward error correction)
- GBT Wide-bus mode: 4.48 Gbps payload
- FULL mode: 9.6 Gbps link speed in 8B/10B

- Scalable architecture
- Routing of event data, detector control, configuration, calibration, monitoring
- E-links configuration configurable: 2/4/8/16 bit
- Detector independent
- TTC (Timing, Trigger and Control) distribution is integrated
- IP blocks are provided.
  - PCIe DMA core: Wupper.
  - Optimized GBT-FPGA core.
  - FULL mode examples for front-end.
- Software:
  - Low-level tools: PCIe driver
  - Control: firmware housekeeping, hardware monitoring and tools to control and monitor the Front-Ends.
  - Testing: DMA & throughput testing. Long time continuous data streaming to the disk, and checking.

The projects using FELIX

- ATLAS Phase-I sub-system
  - Liquid Argon Calorimeter
    - LTDB (LAr Trigger Digitizer Board)
    - LDPB (LAr Digital Processing Blade)
  - Level-1 calorimeter trigger
    - gFEX (Global Feature Extractor)
    - ROD, Hub for eFEX (Electron Feature Extractor) and jFEX (Jet Feature Extractor)
    - TREX (Tile Rear Extension) modules
  - New small wheel of the muon spectrometer
    - sTGC (Small-strip Thin Gap Chamber) and MicroMegas (Micro Mesh Gaseous Structure) detector for muon tracking
- ATLAS Phase-II related projects
  - Test system for Tile Calorimeter
  - CaRIBOu (Control and Readout ITk Board)
    - Test system for Phase-II ITk HV-CMOS pixel sensor R&D
- ProtoDUNE: FELIX is used to readout 1 of the 6 APAs (Anode Plane Assembly).

Total links of FELIX for ATLAS Phase-I

Phase-I System

		Phase-I
MM		512
STGC	p&wFE	256
	sFE	256
RIM		64
Trigger Processor		64
Total		1152

		Phase-I
LTDB	downlink	620
	uplink	620
LDPB	downlink	31
	uplink	155
Total	downlink	651
	uplink	775

		Phase-I
eFEX	downlink	2/4
	uplink	48/96
jFEX	downlink	1/2
	uplink	24/48
gFEX	downlink	4
	uplink	12
TREX	downlink	32
	uplink	32
TOPO	downlink	1
	uplink	24
Total	downlink	40/43
	uplink	140/212

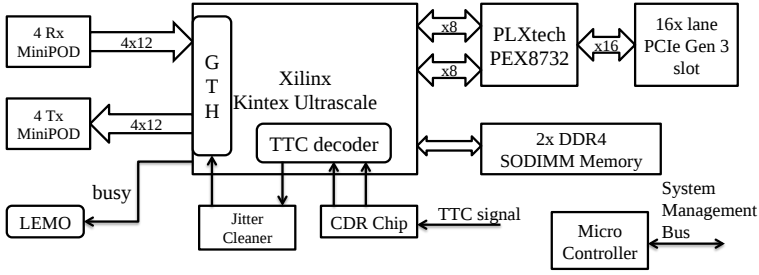
- **GBT @ 4.8 Gbps except full mode @ 9.6 Gbps for LDPB and L1CALO uplinks**

- for GBT links with GBTx (or FPGA) based Front-Ends: 24 links per card.
- for Front-End requires only slow-control and clock distribution like LTDB: ~40+ links per card.
- for FULL mode link with FPGA based on Front-Ends:
  - 12+ Down link: 4.8 Gb/s GBT link.
  - 12 Up link: 9.6 Gb/s FULL mode link, limited by the PCIe bandwidth.

Outline

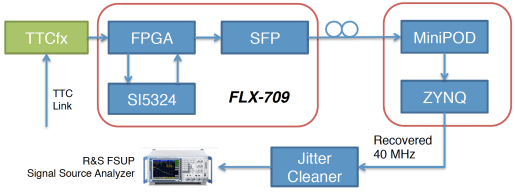
1. The use of BNL-711
2. Design of the BNL-711 PCIe Card
3. Plan for Version 2: FELIX Pre-Production

Design of BNL-711 card



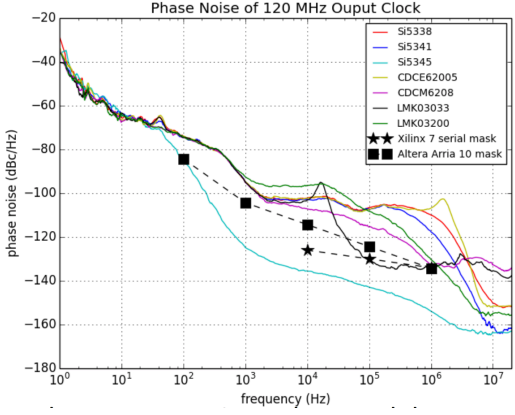
- Basic functions:
  - PCIe Gen 3 $\times$16 lane
  - 48 bidirectional optical links up to 14 Gb/s
  - 2$\times$ DDR4 SODIMM connectors: capacity up to 16 GB
  - with circuits to interface ATLAS TTC (Timing, Trigger and Control) system
  - with on-board jitter cleaner SI5345 and LMK03200: support 0-delay mode
- Micro-controller to support FPGA reprogramming, and firmware version control and update
  - important for detector operation & maintenance

Selection of the clock jitter cleaner



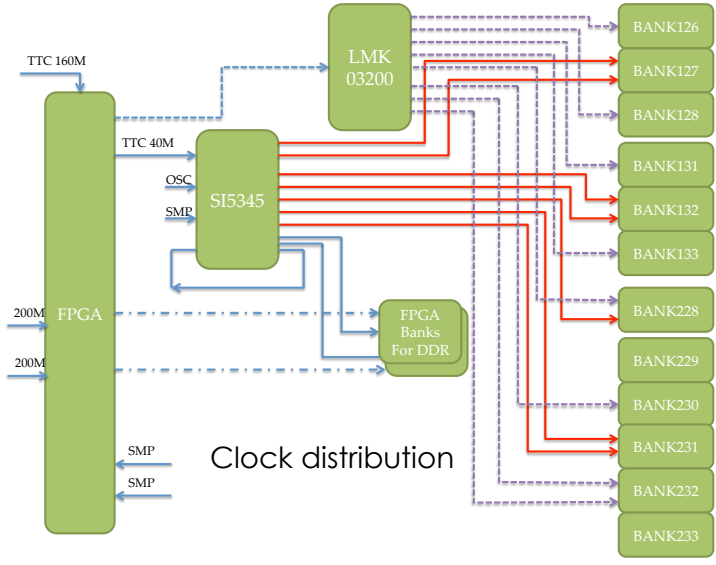
Device	SI5338	SI5345	SI5341
Jitter (ps)	8.58	0.09	6.39
Device	CDCM6208	LMK03200	LMK03033
Jitter (ps)	2.06	5.91	2.74
Device	CDCE62005		
Jitter (ps)	8.61		

The jitter from 10 kHz to 1 MHz



- This survey was originally done to choose a clock device for a FrontEnd board. It also provided input for the selection of jitter cleaner for BNL-711.
- SI5345 showed the best performance.
  - it supports 0-delay mode.
  - 10 outputs.
  - meets requirement for transceivers in both of 7 series FPGA and Ultrascale FPGA.

Clock distribution for the BNL-711 V1P5

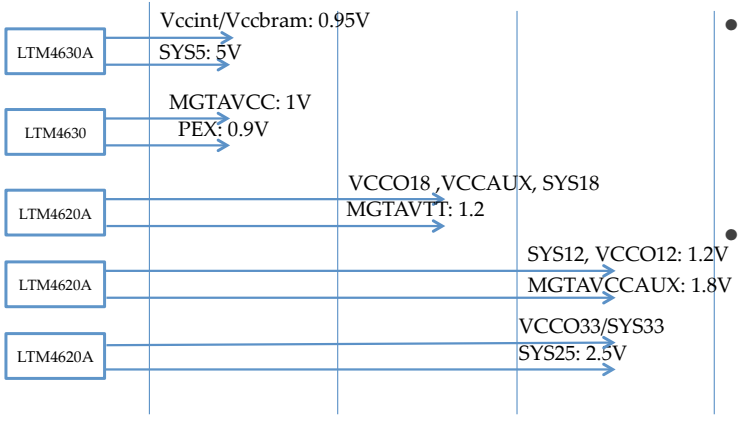


- LMK03200 is a backup for SI5345. Each bank can use reference clocks from both of Si5345 & LMK03200.

Power supply for the BNL-711 V1P5

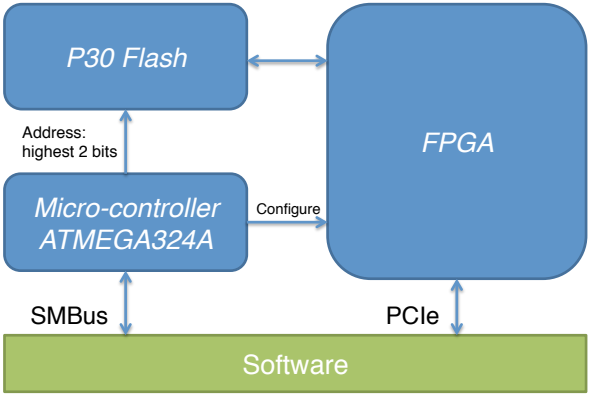
Three stages of power on sequence.

LTM4630A: 18/36A, in: 4.5-15, out: 0.6-5.3
LTM4630: 18/36A, in: 4.5-15, out: 0.6-1.8
LTM4620A: 13/26A, in: 4.5-16, out: 0.6-5.3



- Will use LTM4630A for all the 5× DC-DC power modules.
  - same price.
  - pin compatible.
  - new generation.
- Worst case power dissipation: ∼64W
  - FPGA: about 37 W.
  - others: about 27 W.

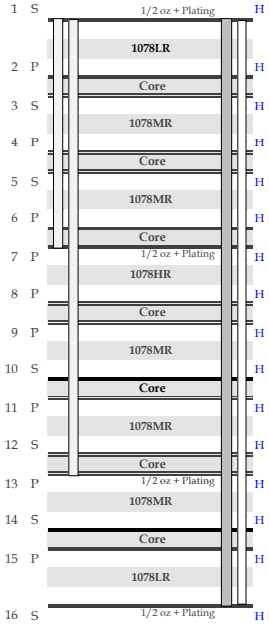
The use of micro-controller



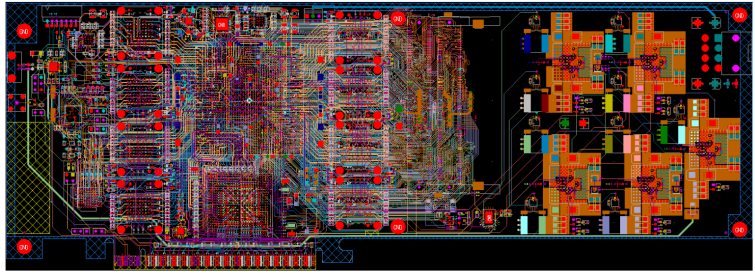
- The flash can store 4 different bit files. It is selected by software via the micro-controller. A golden version can be saved in one of them.
- The software can reprogram the FPGA via micro-controller.
- The FPGA firmware can receive bit file from software via PCIe interface, and update the chosen flash partition.

- *This design is originally from C-RORC (ATLAS RobinNP) board.*
- The program running in the micro-controller and script to control the FPGA programming are modified from the design by Heiko Engel of ALICE group.
- The flash programming via PCIe is being tested.

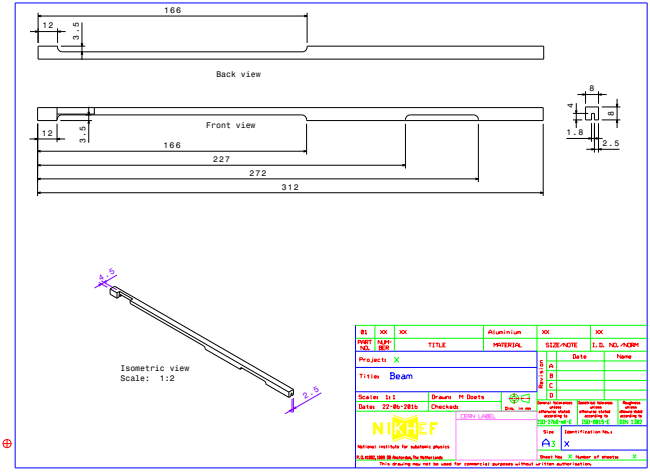
The PCB stackup



- The layout is complicated:
  - board thickness (1.57 mm) limits the number of layers.
  - board height requirement causes the traces to be very dense.
  - special impedance requirement for DDR4.
- Two kinds of blind vias are used. 3 sequential laminations are needed when producing the PCB.
  - one (layer 1-6) is for the MiniPODs.
  - one (layer 1-12) is for DDR4 traces.

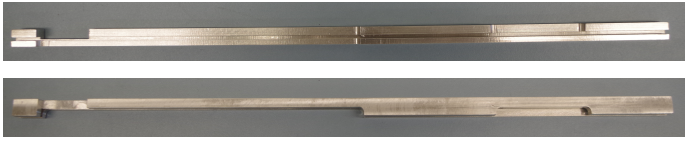


The mechanical design for the BNL-711

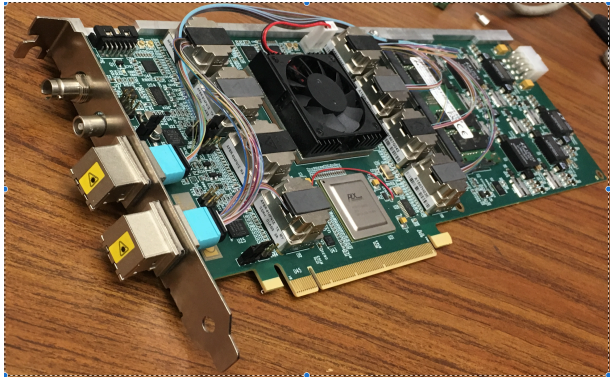


- Form factor of the card:
  - thickness: 1.57mm (62 mils).
  - height: 111.15 mm (4.376 inches)
  - length: 312.00 mm (12.283 inches)

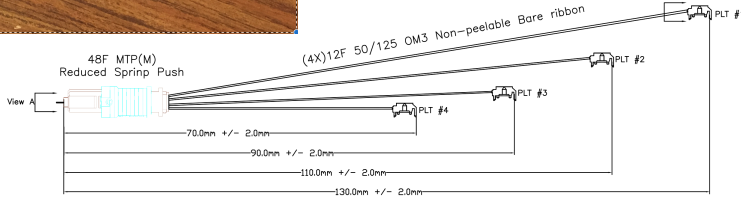
- The stiffener bar for this long board.
  - designed by Nikhef
  - machined at BNL.
  - will be glued on the board.



The front bracket and fiber



- The front panel:
  - TTC fiber
  - LEMO connecter for BUSY signal.
  - two MTP couplers, in each:
    - 24 Tx channels
    - 24 Rx channels



- Two harnesses are ordered, each has 4× 12-channel fibers with different lengths to match the fiber routing.

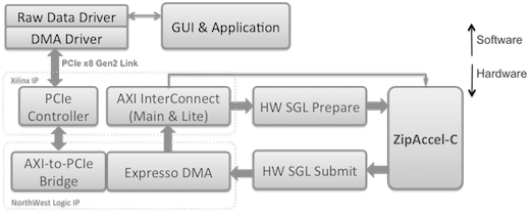
The FPGA selection

Package	F1924	B2104	F1924/B2104	B2104	B2104	B2104
FPGA	KU085	KU095	KU115	VU125	VU160	VU190
Logic Cells	1088k	1176k	1451k	1566k	2026k	2350k
LUT	497k	537k	663k	716k	926k	1074k
Distributed RAM	13.4 Mb	4.7 Mb	18.3 Mb	9.7Mb	12.7 Mb	14.5 Mb
Block RAM	56.9 Mb	59.1 Mb	75.9 Mb	88.6 Mb	115.2 Mb	132.9 Mb
PCIe Gen3 x8	4	4	4	4	4	6
GTH Transceiver	56	32 (16.3G)	64	40	40	40
GTY Transceiver	0	32(16.3G)	0	36 (30.5G)	36 (30.5G)	36 (30.5G)
Price		6059	7554		22310	26310

- Here F1924 is FLVF1924. B2104 is FLVB2104 for KU115 and VU125. B2104 is FLGB2104 for VU160 and VU190.
- The average cost per fiber link for Kintex Ultrascale is the lowest.
- Current package is FLVF1924, default FPGA is XCKU115-2. For KU085 with same package, besides 16 PCIe lanes, 40 links are available. For V1.5, if using KU085, the DDR4B and one pair of SMP test points can't be used.

Available resources for sPHENIX TPC DAM

Package	F1924	F1924
FPGA	KU085	KU115
Logic Cells	1176k	1451k
LUT	497k	663k
Distributed RAM	13.4 Mb	18.3 Mb
Block RAM	56.9 Mb	75.9 Mb
PCIe Gen3 x8	4	4
GTH Transceiver	56	64
GTY Transceiver	0	0
Price		7554



- Based on the discussion with Jin, some important requirement of sPHENIX TPC can be met:
  - Buffer 20 us of data with RAM inside FPGA: 80Gb/s×20us=1.6 Mb. This is much less than resource of KU085 & KU115.
  - The data compression in FPGA firmware: the ZipAccel-C IP found by Jin can handle 100Gbps input with only 15k LUTs. This IP supports Deflate (RFC-1951), ZLIB (RFC-1950) & GZIP (RFC-1952). Links:
    - https://www.xilinx.com/products/intellectual-property/1-7aisy9.html
    - http://www.cast-inc.com/ip-cores/data/zipaccel-c/index.html

Some test results and parameters of BNL-711

- Throughput of PCIe: ~101.7 Gb/s.
- DDR4 speed: up to 2.1 GT/s.
- The 48 links: BER < 1E-15 for 12.8 Gb/s.

Table 35: **Clock Buffers Switching Characteristics**

Symbol	Description	Speed Grade and V_{CCINT} Operating Voltages				Units
		1.0V	0.95V		0.90V	
		-3	-2	-1/-1L	-1L	
Global Clock Switching Characteristics (Including BUFGCTRL)						
F_{MAX}	Maximum frequency of a global clock tree (BUFG)	850	725	630	630	MHz
Global Clock Buffer with Input Divide Capability (BUFGCE_DIV)						
F_{MAX}	Maximum frequency of a global clock buffer with input divide capability (BUFGCE_DIV)	850	725	630	630	MHz
Global Clock Buffer with Clock Enable (BUFGCE)						
F_{MAX}	Maximum frequency of a global clock buffer with clock enable (BUFGCE)	850	725	630	630	MHz
Leaf Clock Buffer with Clock Enable (BUFCE_LEAF)						
F_{MAX}	Maximum frequency of a leaf clock buffer with clock enable (BUFCE_LEAF)	850	725	630	630	MHz
GTH/GTY Clock Buffer with Clock Enable and Clock Input Divide Capability (BUFG_GT)						
F_{MAX}	Maximum frequency of a serial transceiver clock buffer with clock enable and clock input divide capability	512	512	512	512	MHz

Some parameters of BNL-711

Table 32: **Block RAM and FIFO Switching Characteristics**

Symbol	Description	Speed Grade and V_{CCINT} Operating Voltages				Units
		1.0V	0.95V		0.90V	
		-3	-2	-1/-1L	-1L	
Maximum Frequency						
$F_{MAX\_WF\_NC}$	Block RAM (Write First and No Change modes)	660	585	525	525	MHz
F_{MAX_RF}	Block RAM (Read First mode)	575	510	460	400	MHz
F_{MAX_FIFO}	FIFO in all modes without ECC	660	585	525	525	MHz
F_{MAX_ECC}	Block RAM and FIFO in ECC configuration without PIPELINE	530	450	390	390	MHz
	Block RAM and FIFO in ECC configuration with PIPELINE and Block RAM in Write First or No Change mode.	660	585	525	525	MHz
	Block RAM in ECC configuration in Read First mode with PIPELINE	575	510	460	400	MHz
$F_{MAX\_ADDREN\_RDADDRCHANGE}$	Block RAM with address enable and read address change compare turned on.	575	510	460	400	MHz
$T_{PW\_WF\_NC}$(1)	Block RAM in WRITE_FIRST and NO_CHANGE modes and FIFO. Clock High/Low pulse width	758	855	952	952	ps, Min
T_{PW_RF}(1)	Block RAM in READ_FIRST modes. Clock High/Low pulse width	870	980	1087	1250	ps, Min

Table 34: **DSP48 Slice Switching Characteristics**

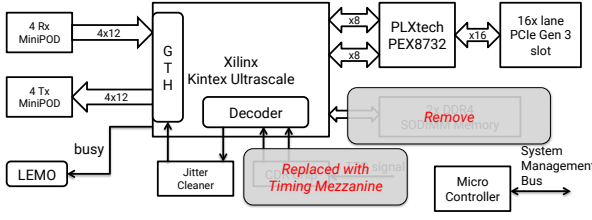
Symbol	Description	Speed Grade and V_{CCINT} Operating Voltages				Units
		1.0V	0.95V		0.90V	
		-3	-2	-1/-1L	-1L	
Maximum Frequency						
F_{MAX}	With all registers used	741	661	594	594	MHz
F_{MAX_PATDET}	With pattern detector	687	581	512	512	MHz
$F_{MAX\_MULT\_NOMREG}$	Two register multiply without MREG	462	429	361	361	MHz
$F_{MAX\_MULT\_NOMREG\_PATDET}$	Two register multiply without MREG with pattern detect	428	387	326	326	MHz
$F_{MAX\_PREADD\_NOADREG}$	Without ADREG	468	429	358	358	MHz
$F_{MAX\_NOPIPELINEREG}$	Without pipeline registers (MREG, ADREG)	335	312	260	260	MHz
$F_{MAX\_NOPIPELINEREG\_PATDET}$	Without pipeline registers (MREG, ADREG) with pattern detect	316	286	238	238	MHz

Outline

1. The use of BNL-711
2. Design of the BNL-711 PCIe Card
3. Plan for Version 2: FELIX Pre-Production

Plan for version 2

- Remove DDR4 modules, give more space for other design, also make the board to be shorter.
- Change the mapping PCIe endpoint mapping, put one in each of the two SLR (Super Logic Region).
- Remove TTC related circuits, add a connector for timing mezzanine card, to support different timing systems:
  - TTC
  - TTC-PON
  - White Rabit
  - sPHENIX TPC (We can help the design of mezzanine).



Schedule for version 2

- Schedule:
  - Two V1.5 cards have been tested. Two V1.5 is being assembled.
  - 02/2017: V1.5 will be fully tested by FELIX group, firmware and software will be ready in 03/2017.
  - 01/2017: design of the pre-production board V2.0 starts.
  - 01/2017 - 02/2017: Schematics design
  - 03/2017 - 05/2017: Layout design
  - 06/2017 - 07/2017: Fabrication and assembly
  - 08/2017: Initial evaluation test
  - 09/2017 - 10/2017: Assembly and test of more V2.0 BNL-711
  - 10/2017: V2.0 BNL-711 is available for firmware development

What can we provide for sPHENIX TPC DAM

- Boards after the initial evaluation test.
- Support of the reusable firmware & software development from current designs.
- Help design the mezzanine card for the special timing system.
- Additional hardware functions if compatible with the FELIX project.

Thanks!